

ACOUSTIC SIGNAL ENCODING METHOD AND DECODING METHOD

Publication number: JP8263096 (A)

Publication date: 1996-10-11

Inventor(s): JIN AKIO; MORIYA TAKEHIRO; MIKI SATOSHI

Applicant(s): NIPPON TELEGRAPH & TELEPHONE

Classification:

- international: **G10L19/02; G10L19/00; G10L19/04; G10L19/08; G10L19/00;**
(IPC1-7): G10L7/04; G10L9/14; G10L9/18

- European:

Application number: JP19950065622 19950324

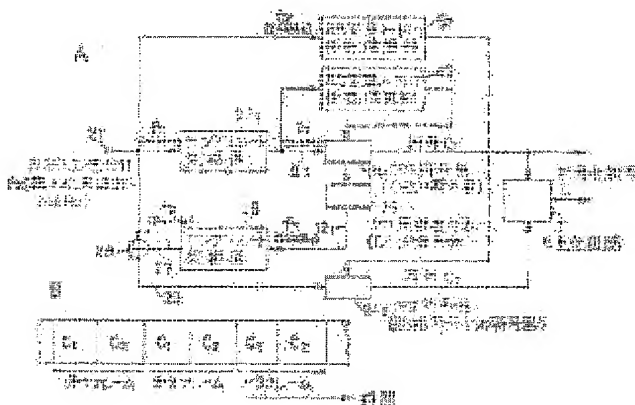
Priority number(s): JP19950065622 19950324

Also published as:

 JP3139602 (B2)

Abstract of JP 8263096 (A)

PURPOSE: To encode a sound at a high compression rate and to encode a musical tone with high quality by using a CELP system and a conversion coding system. **CONSTITUTION:** An input signal 11 of a sampling frequency $f_s=24\text{kHz}$ is made a low band signal of $f_s=16\text{kHz}$ by a converter 221, and it is encoded by a CELP coder 241, and a resultant code C1 is outputted, and the code C1 is decoded by a decoder 251, and the decoded signal is made the signal of $f_s=24\text{kHz}$ by a converter 26, and it is subtracted from the input signal 11, and a high band signal and a quantization error signal are coded by a conversion coding coder 242, and the code C2 is outputted. Only the code C1, or both of C1 and C2 are decoded to be used.



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平8-263096

(43)公開日 平成8年(1996)10月11日

(51)Int.Cl. ⁸	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L	7/04		G 1 0 L	G
	9/14			G
				J
	9/18			E
				C

審査請求 未請求 請求項の数6 O L (全 11 頁)

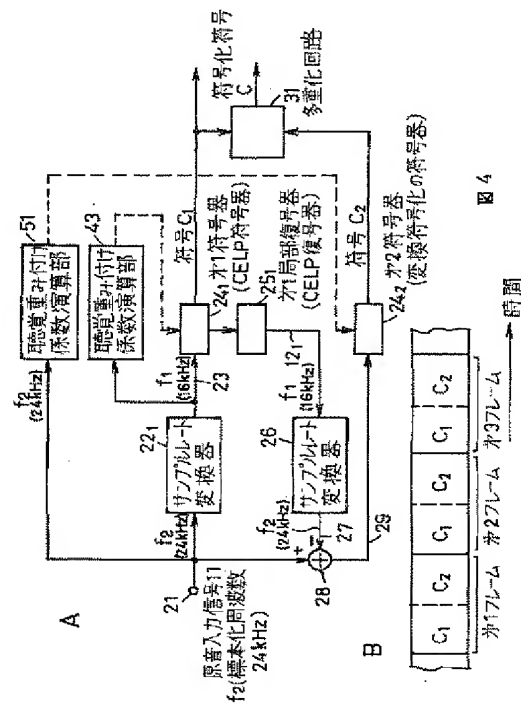
(21)出願番号	特願平7-65622	(71)出願人	000004226 日本電信電話株式会社 東京都新宿区西新宿三丁目19番2号
(22)出願日	平成7年(1995)3月24日	(72)発明者	神 明夫 東京都千代田区内幸町1丁目1番6号 日本電信電話株式会社内
		(72)発明者	守谷 健弘 東京都千代田区内幸町1丁目1番6号 日本電信電話株式会社内
		(72)発明者	三樹 聡 東京都千代田区内幸町1丁目1番6号 日本電信電話株式会社内
		(74)代理人	弁理士 草野 卓

(54) 【発明の名称】 音響信号符号化方法及び復号化方法

(57) 【要約】

【目的】 CELP方式と、変換符号化方式とを用い、音声を高い圧縮率で符号化し、楽音を高い品質で符号化する。

【構成】 標準化周波数 $f_s = 24 \text{ kHz}$ の入力信号 11 を変換器 22₁ で $f_s 16 \text{ kHz}$ の低域信号とし、これを CELP 符号器 24₁ で符号化して符号 C₁ を出力し、その符号 C₁ を復号器 25₁ で復号し、その復号信号を変換器 26 で $f_s = 24 \text{ kHz}$ の信号とし、これを入力信号 11 から差引き、高域信号と量子化誤差信号とを変換符号化符号器 24₂ で符号化して符号 C₂ を出力する。符号 C₁ のみ、又は C₁ と C₂ の両方を復号して利用する。



【特許請求の範囲】

【請求項1】 楽音や音声などの最高周波数が f_n の音響入力信号を周波数 f_1, f_2, \dots, f_{n-1} ($f_1 < f_2 < \dots < f_{n-1} < f_n$)の n 個の区分(n は2以上の整数)に分割して符号化する符号化方法において、

上記入力信号から周波数が f_1 以下の第1帯域信号を選出する第1帯域選択過程と、

上記第1帯域信号を第1符号化方法で符号化して第1符号を出力する第1符号化過程と、

第 $i-1$ 以下の各符号($i=2, 3, \dots, n$)から周波数が f_{i-1} 以下の第 $i-1$ 復号信号を得る第 $i-1$ 復号化過程と、

上記入力信号から周波数 f_i 以下の第 i 帯域信号を選出する第 i 選択過程と、

上記第 i 帯域信号から上記第 $i-1$ 復号信号を差し引いて第 i 差信号を得る第 i 差過程と、

上記第 i 差信号を第 i 符号化方法で符号化して第 i 符号を出力する第 i 符号化過程と、

を有する音響信号符号化方法。

【請求項2】 上記第 $i-1$ 復号化過程は上記第 $i-1$ 符号を復号する過程と、その復号された信号と第 $i-2$ 復号信号とを加算する過程と、その加算された信号を標本化周波数が $2f_1$ の信号に変換して上記第 $i-1$ 復号信号を得る過程と、

を有することを特徴とする請求項1記載の音響信号符号化方法。

【請求項3】 楽音、音声などの最高周波数が f_n の音響入力信号を、周波数 f_1, f_2, \dots, f_{n-1} ($f_1 < f_2 < \dots < f_{n-1} < f_n$) ($n=2$ 以上の整数)で区分してそれぞれを符号化する符号化方法において、

上記入力信号より標本化周波数が $2f_1$ の第1帯域信号を得る第1帯域選択過程と、

上記第1帯域信号を第1符号化法により符号化して第1符号を出力する第1符号化過程と、

上記 $i-1$ 符号化過程($i=2, 3, \dots, n$)の符号誤差として第 $i-1$ 誤差信号を得る第 $i-1$ 誤差取出し過程と、

上記第 $i-1$ 誤差信号を標本化周波数が $2f_1$ の第 $i-1$ 変換誤差信号に変換する第 $i-1$ 変換過程と、

上記入力音響信号より周波数帯域が $f_{i-1} \sim f_i$ 、標本化周波数が $2f_1$ の第 i 帯域信号を得る第 i 帯域選出過程と、

上記第 $i-1$ 変換誤差信号と上記第 i 帯域信号とを加算して第 i 加算信号を得る第 i 加算過程と、

上記第 i 加算信号を第 i 符号化法により符号化して第 i 符号を出力する第 i 符号化過程と、

を有する音響信号符号化方法。

【請求項4】 上記第1符号化法は符号駆動線形予測符号化法であり、上記第 n 符号化法は変換符号化法である

ことを特徴とする請求項1乃至3の何れかに記載の音響信号符号化方法。

【請求項5】 上記音響入力信号中の周波数 f_i 以下のほぼ全域の成分のスペクトル包絡を重みの基準として、上記第 i 符号化過程において心理聴覚重み付け量子化を行うことを特徴とする請求項1乃至4の何れかに記載の音響信号符号化方法。

【請求項6】 入力符号を第1乃至第 n 符号(n は2以上の整数)に分離する分離過程と、

上記第1符号を復号して、標本化周波数 $2f_1$ の第1復号信号を第1復号化出力として出力する第1復号過程と、

上記第 $i-1$ 復号化出力($i=2, 3, \dots, n$)を標本化周波数が $2f_i$ の第 $i-1$ 変換復号化出力に変換する第 $i-1$ 変換過程と、

上記第 i 符号を復号して標本化周波数 $2f_i$ の第 i 復号信号を得る第 i 復号過程と、

上記第 i 復号信号と上記第 $i-1$ 変換復号化出力とを加算して第 i 復号化出力を出力する第 i 加算過程と、

を有する音響信号復号化方法。

【発明の詳細な説明】

【0001】

【産業上の利用分野】この発明は、楽音や音声などの音響信号を周波数領域で帯域分割して階層符号化する符号化方法及びその復号化方法に関する。

【0002】

【従来の技術】音響信号を周波数領域で帯域分割して符号化する方法として、サブバンド符号化方法がある。サブバンド符号化方法はQMF(Quadrature Mirror Filter)を用いて入力信号を複数の周波数帯域に分割し、その各帯域に適切なビット割り当てを行いつつ各帯域を独立に符号化する。

【0003】現在、楽音及び音声などの音響信号の符号化方法は使用目的、復号品質、符号化速度などに応じて多種多様な方法があるが、1つの音響信号に対して複数の符号化方法を得ることなく1つの符号化方法でのみ符号化するのが普通である。しかし、例えば図1Aに示すように音響信号11を周波数軸上で低域側から3つのサブバンド SB_1, SB_2, SB_3 に分割して階層化し、図2に示すようにその下位層(階層1)であるサブバンド SB は符号化品質は低い、すなわち復号再生音の周波数帯域が狭く、量子化誤差も大きい符号化方法、例えば符号駆動線形の予測符号化法:CELPにより高圧縮率で符号化し、逆に上位層(階層3)であるサブバンド SB_3 の符号化は符号化品質が高く、すなわち復号再生音の周波数帯域が広く、量子化誤差が小さい符号化方法

(例えば離散コサイン変換符号化方法などの変換符号化法で低圧縮率で符号化し、中位層(階層2)であるサブバンド SB_2 に対しては下位層の符号化方法と、上位層の符号化方法との中間の符号化方法とし、利用者の要求

に応じて階層1のみを符号化送出し、あるいは階層1と2を符号化送出し、又は全ての階層を符号化送出するという符号化方法も、考えられる。

【0004】あるいは前述のように3つに階層符号化された各種の楽音又は音声信号を例えばデータベースとして設け、利用者からのそのデータベースをアクセスし、所望の楽音信号を受け取り、その利用者の復号器に応じて、階層1の符号のみを復号して狭帯域かつ量子化誤差の大きい低品質の再生音を得、あるいは階層1及び2の符号を復号、又は階層1、2、3の全ての符号を復号して広帯域かつ量子化誤差の小さい高品質な再生音を得ることが考えられる。

【0005】又は、例えば、音声が支配的な広帯域の音響信号を2階層に分けて符号化し、その下位層符号のみを復号すれば主に音声的な性質を有する音響信号をきれいに復号し、下位層と上位層との両符号を復号すれば、更に、非音声的な性質を有する音響信号も含めた信号の復号ができる、ということが考えられる。またこれらの場合において、下位層符号のみを受け取り、その際の伝送路の利用時間を短かくしまたは、伝送容量の小さな伝送路を使用し、かつ実時間で復号したり、長い時間かけて上位層符号をも受け取り、一度蓄積した後、改めて再生復号することにより高品質の復号信号を得ることもできる。

【0006】あるいは、これらの場合において、下位・上位層の全ての符号を一度蓄積した後、下位層符号のみを、小型かつ経済的な遅延時間の小さい復号器により実時間で復号したり、高品質な音を再生したい時には、上位層符号をも含めて、大型かつ遅延時間の大きな復号器により、時間をかけて復号し、その後で一度に再生することもできる。

【0007】前述のように復号品質や符号化圧縮率に選択性をもたせる符号化方法はスケーラブルな階層符号化方法と称せられる。スケーラブルな階層符号化方法としては図1Aに示したサブバンド符号化方法が考えられる。すなわち符号化方法1によってサブバンド SB_1 の周波数帯域を符号化し、同様にして帯域 SB_2 、 SB_3 を各々独立した符号化方法2、3により符号化を実行する。図1Bに示すように、復号化の際には、例えば、広帯域の復号音を必要としない時には、サブバンド SB_1 の符号のみを符号化方法1の復号器により復号化して、サブバンド SB_1 の帯域のみの音の復号信号 12_1 を得、また広帯域復号音を必要とする場合はサブバンド SB_1 、 SB_2 、 SB_3 の各符号をそれぞれ符号化方法1、2、3と対応した復号量により復号して復号信号 12_1 、 12_2 、 12_3 を得てこれらの合成信号 12 を出力する。

【0008】

【発明が解決しようとする課題】しかし、このようなサブバンド符号化方法による階層符号化では、各帯域（す

なわち各層）に発声する量子化誤差、すなわち符号器の入力信号とその局部復号器の出力信号、つまり伝送路などの影響を受けていない復号信号との誤差が図1Cに示すように各帯域 SB_1 、 SB_2 、 SB_3 にそれぞれ量子化誤差 13_1 、 13_2 、 13_3 として保存され、よって全周波数帯域の復号信号 12 には各帯域毎に独立に歪みや雑音が発生してしまう。従って、全帯域を復号する場合（すなわち上位層までの復号化）でさえも、下位層の大きな量子化誤差 13_1 も、そのまま発生するため、高品質のものは得られない。広帯域復号信号を高品質に得るには各サブバンド SB_1 、 SB_2 、 SB_3 での各符号化圧縮率を小さくしなければ、量子化雑音を低減させることができない。従ってこのような階層符号化方法では、スケーラブルな符号化を実現できない。

【0009】従来のサブバンド符号化方法によるスケーラブルな符号化ができないことを図3を参照して更に具体的に説明する。即ち原音響信号 11 の帯域を2分割し、第1階層（低域領域）をCELP方式で符号化し、第2階層（高域領域）を変換符号化方法により符号化している。第1階層では、音声の圧縮効率の高いCELP符号化が実行されているため、その局部復号信号 12 、（図3B）の量子化誤差信号 13_1 は図3Cに示すように比較的大きい。一方第2階層では様々な波形に対して符号化可能な変換符号化が実行されているため、その局部復号信号 12_2 は図3Bに示すように原音信号 11 に近く、量子化誤差信号 13_2 は図3Cに示すように小さい。しかし第1階層の符号及び第2階層の符号をそれぞれ復号して広域復号信号を得ても、図3Dに示すようにその復号信号の量子化誤差の低域部分 14_1 は第1階層の量子化誤差 13_1 と変わらない。すなわち、第2階層までの復号品質は低周波数の帯域においてCELP符号化方法の符号化性能に依存してしまう。よって、サブバンド符号化方法で階層符号化を行い高品質な符号化品質を実現するためには、各階層全てを圧縮率が小さいか、または演算量の大きな高品質符号化方法によって符号化しなければならない。

【0010】この発明の目的は、下位層での符号化を高圧縮率、低復号品質とし、しかも上位層までの復号信号に下位層の低復号品質の影響を受けない高品質のものを得ることができるスケーラブルな符号化方法及びその復号化方法を提供することにある。

【0011】

【課題を解決するための手段】請求項1の発明によれば、楽音や音声などの最高周波数が f_n の音響入力信号を周波数 f_1 、 f_2 、……、 f_{n-1} （ $f_1 < f_2 < \dots < f_{n-1} < f_n$ ）の n 個の区分（ n は2以上の整数）に分割して符号化する符号化方法において、入力信号から周波数が f_1 以下の第1帯域信号を選出し、その第1帯域信号を第1符号化方法で符号化して第1符号を出力し、第 $i-1$ 以下の各符号（ $i=2, 3, \dots$ 、

n) から周波数が f_{i-1} 以下の第 $i-1$ 復号信号を得、上記入力信号から周波数 f_i 以下の第 i 帯域信号を選出し、その第 i 帯域信号から上記第 $i-1$ 復号信号を差し引いて第 i 差信号を得、その第 i 差信号を第 i 符号化方法で符号化して第 i 符号を出力する。

【0012】第 $i-1$ 復号信号は例えば、第 $i-1$ 符号を復号化した信号と、第 $i-2$ 復号信号とを加算し、その加算信号を標本化周波数が $2f_i$ の信号に変換して得る。請求項3の発明の符号化方法によれば楽音、音声などの最高周波数が f_n の音響入力信号を、周波数 f_1, f_2, \dots, f_{n-1} ($f_1 < f_2 < \dots < f_{n-1} < f_n$) ($n=2$ 以上の整数) で区分してそれぞれ符号化する符号化方法において、上記音響入力信号より標本化周波数が $2f_i$ の第1帯域音響信号を得、その第1帯域信号を第1符号化法により符号化して第1符号を出力し、その第1符号の符号化誤差を第 $i-1$ 誤差信号を得 ($i=2, 3, \dots, n$)、その第 $i-1$ 誤差信号を標本化周波数が $2f_i$ の第 $i-1$ 変換誤差信号に変換し、上記音響入力信号より周波数帯域が $f_{i-1} \sim f_i$ 、標本化周波数が $2f_i$ の第 i 帯域信号を得、上記第 $i-1$ 変換誤差信号と上記第 i 帯域信号とを加算して第 i 加算信号を得、その第 i 加算信号と第 i 符号化法により符号化して第 i 符号を出力する。

【0013】請求項4の発明のよれば、請求項1乃至3の何れかの発明において上記第1符号化法は符号駆動線形予測符号化法であり、上記第 n 符号化法は変換符号化法である。請求項5の発明では請求項1乃至4の何れかの発明において、上記音響入力信号中の周波数 f_i 以下のほぼ全域の成分のスペクトル包絡を重みの基準として、上記第 i 符号の符号化過程において心理聴覚重み付け量子化を行う。

【0014】請求項6の発明の復号化方法によれば、入力符号を第1乃至第 n 符号 (n は2以上の整数) に分離し、上記第1符号を復号して、標本化周波数 $2f_1$ の第1復号信号を出力し、上記第 $i-1$ 復号信号 ($i=2, 3, \dots, n$) を標本化周波数が $2f_i$ の第 $i-1$ 変換復号信号に変換し、上記第 i 符号を復号して標本化周波数 $2f_i$ の第 i 復号信号を得、その第 i 復号信号と上記第 $i-1$ 変換復号信号とを加算して第 i 加算信号を出力する。

【0015】

【実施例】図4Aに請求項1の発明の符号化方法の実施例を適用した符号化器の例を示す。この例では原音信号を2つの周波数帯域に分けて符号化、つまり2階層符号化する場合である。入力端子21からの原音入力信号11は標本化周波数が24kHz、つまり最高周波数 f_2 が12kHzのデジタル信号であり、この入力信号は第1帯域選択手段としてのサンプルレート変換器22₁で標本化周波数が16kHzの信号に変換されて第1帯域信号23が取出される。このサンプルレート変換はいわ

ゆるダウンサンプリングであり、例えば変換標本化周波数比に応じた間隔でサンプルが除去された後、デジタル低域通過フィルタを通して実行される。このサンプルレート変換器22₁よりの周波数が $f_1=8$ kHz以下の第1帯域信号23が取出され、この第1帯域信号23は第1符号化方法による第1符号器24₁で符号化される。この例では第1符号器24₁としてCELP(符号駆動線形予測符号化)符号方法により符号化する。この符号化の結果である第1符号 C_1 が出力される。

【0016】この実施例では局部復号器25₁で復号され、周波数が f_1 以下の第1復号信号12₁が得られ、その復号信号12₁は第1サンプルレート変換器26₁で標本化周波数が24kHzの変換復号信号27に変換される。このサンプルレート変換器26₁はいわゆるアップサンプリングを行うものであり、例えば、変換周波数比に応じた間隔でゼロサンプルを加えた後デジタル低域通過フィルタに通せばよい。差回路28で入力信号11からこの変換復号信号27が差し引かれ、その差信号29が第2符号化方法による第2符号器24₂で符号化される。この実施例では第2符号器24₂で変形離散コサイン変換などの変換符号化(Transform coding)により符号化される。この符号化結果の第2符号 C_2 は出力される。第1符号 C_1 と第2符号 C_2 は多重化回路31で、例えば図4bに示すように符号化フレームごとに時分割的に多重化され、符号化符号 C として出力される。利用者の要求によっては第1符号 C_1 のみを出力してもよい。

【0017】標本化周波数24kHzの原音入力信号11の周波数スペクトルは例えば図5Aに示され、この信号11中の8kHz以下の信号が標本化周波数16kHzの信号23(図5B)として下位層の第1符号器24₁に入力され、高い圧縮効率で符号化される。その符号化符号 C_1 の局部復号器25₁により復号された復号信号12₁は図5Bに示すように、下位層入力信号23に対しては少なからず量子化誤差13₁が図5Cに示すように生じる。差回路28からこの誤差信号13₁と、原音入力信号11の8kHz以上の高域信号33とよりなる信号29が上位層の第2符号器24₂に入力され、あらゆる性質の入力信号を高品質に符号化可能な変換符号化方法により符号化される。

【0018】このようにこの実施例では下位層の符号化符号 C_1 は原音をそれ程忠実には符号化しないが、上位層では下位層の量子化誤差も含めて符号化されるため、後述で明らかにするように、上位層まで復号する場合に、下位層をも高い忠実度で復号再生することが可能となる。つまり下位層では高い圧縮効率で符号化し、しかも上位層をも復号する場合は、高品質の復号信号を得ることができる。

【0019】特に前記実施例では下位層の符号化にCELP方式を用いているため符号化対象が音声の場合、下

位層の第1符号 C_1 のみを復号しても比較的良好な品質が得られ、また演算量が少なく、実時間処理が容易である。第1、第2符号 C_1 、 C_2 を復号して、符号化対象が楽音であっても、上位層の変換符号の復号により、かつ下位層のCELP符号の符号化誤差の補償により、広帯域にわたり、品質の高い復号信号が得られる。

【0020】符号化を行う場合に、人間の心理聴覚、例えば大きいレベルのスペクトルによるマスキング特性などを考慮して、心理聴覚重み付けをして符号化することにより聴覚的に量子化誤差を抑圧した効率的な符号化をすることがよくある。例えば符号器24₁のCELP符号化方法においては図6に示すように、制御部35により指定される周期(ピッチ)のベクトルが適応符号帳36から取出され、また指定された雑音符号帳37から雑音ベクトルが取出され、これらはそれぞれ利得が付与された後、合成されて線形予測合成フィルタ38に励振ベクトルとして入力される。一方図4Aのサンプルレート変換器22₁よりの入力信号は符号化フレーム周期で線形予測分析部39で線形予測分析され、その線形予測係数が量子化部41で量子化され、その量子化線形予測係数に応じて合成フィルタ38のフィルタ係数が設定される。また聴覚重み付け係数演算部43で線形予測係数より求めたスペクトル包絡に基づいて心理聴覚重み付けのためのフィルタ係数を求めて、聴覚重み付けフィルタ42に設置する。サンプルレート変換器22₁よりの入力信号から合成フィルタ38よりの合成信号が差し引かれ、その差信号が聴覚重み付けフィルタ42へ通され、その出力のエネルギーが最小になるように制御部35により適応符号帳36、雑音符号帳37に対する選択が行われる。

【0021】変換符号器24₂の変換符号化方法においては、例えば図7に示すように差回路器28の出力が離散コサイン変換器45で直交コサイン変換されて周波数領域の係数に変換され、そのスペクトル包絡成分が線形予測分析部46で線形予測分析され、これよりスペクトル包絡を得、そのスペクトル包絡で変換器45の出力係数が割算されて正規化され、その平均化された係数が聴覚重み付け部47で聴覚重み付けがなされ、更に量子化部48で例えばベクトル量子化される。聴覚重み付け係数を得るため、この実施例について入力端子21から原音入力信号11が離散コサイン変換器49で直交コサイン変換して、周波数領域に変換され、その変換係数のスペクトル包絡にもとづいて聴覚重み付け係数が係数演算部51で演算されて聴覚重み付け部47に与えられ、正規化係数の対応する成分に対する乗算がなされる。

【0022】つまり、上位層の第2の符号器24₂では図5Cに示すスペクトルの信号29を符号化するが、この信号29のスペクトル包絡にもとづいて聴覚重み付けを行うのではなく、原音入力信号11のスペクトル包絡(図5D)を求め、これに基づいて聴覚重み付け符号化

を行う。次にこの発明の復号化方法の実施例を図8を参照して説明する。この実施例は図4に示した符号化法による符号化符号の復号化に適用した場合である。入力端子55より入力された入力符号に分離回路56で第1符号 C_1 と第2符号 C_2 とに分離され、第1符号 C_1 は第1復号器57₁によりこの例ではCELP復号化方法により最高信号周波数 f_1 (標本化周波数16kHz)の第1復号信号58₁に復号されて下位層(低域)復号化出力63₁として出力される。

【0023】この第1復号化出力58₁はサンプルレート変換器59により最高信号周波数 f_2 (標本化周波数が24kHz)の変換復号信号61₁に変換される。一方分離回路56よりの第2符号 C_2 は第2復号器57₂によりこの例では変換符号復号化がなされ、最高信号周波数 f_2 (標本化周波数が24kHz)の第2復号信号58₂が得られて、この第2復号信号58₂は第1変換復号信号61₁と加算器62₂で加算されて上位層(全帯域)復号化出力63₂として出力される。

【0024】つまり下位層復号化出力63₁としては理想的な場合は図5B中の復号信号12₁が得られる。一方第2復号器57₂の復号信号58₂は理想的には図52Eに示すように、下位層(低域)の量子化誤差信号13₁の復号信号60₁と、高域信号33の復号信号64₂とである。よって加算器62₂よりの復号化出力63₂には低域の復号信号58₁に対し、その量子化誤差13₁と対応する復号信号60₁が加算され、量子化誤差が著しく軽減され、かつ高域復号信号64₂に高い忠実度のものであるから、加算器62から得られる上位層までの復号化出力63₂は原音入力信号11に著しく近く、その量子化誤差信号は例えば図5Fに示すように全帯域にわたり、著しく小さなものとなる。

【0025】次にこの発明の符号化方法を n 階層(n 帯域)分割符号化に適用した例として $n=4$ の場合について図9を参照して説明する。図9において図4Aと対応する部分に同一符号を付けてある。この例では原音入力信号11は最高周波数が $f_n=f_4$ でその標本化周波数が $2f_4$ であり、第1サンプルレート変換器(第1帯域選択手段)22₁で標本化周波数が $2f_1$ (但し $f_1 < f_2 < f_3 < f_4$)の入力信号23₁に変換され、つまり周波数 f_1 以下の第1帯域信号23₁が選出され、その第1帯域信号23₁は第1符号器24₁で符号化され、第1符号 C_1 として出力されると共にその第1符号 C_1 は第1復号器25₁で標本化周波数 $2f_1$ の信号に復号され、その復号信号12₁は第1サンプルレート変換器26₁で標本化周波数が $2f_2$ の第1変換復号信号に変換される。一方入力信号11が第2帯域選択手段としてのサンプルレート変換器22₂で標本化周波数が $2f_2$ の信号に変換されて、周波数 f_2 以下の第2帯域信号23₂が取出される。この第2帯域信号23₂から第1サンプルレート変換器26₁よりの第1変換復号信号

が第2差回路28₂で引算され、その第2差信号29₂が第2符号器24₂で符号化され、第2符号C₂が出力される。

【0026】以下同様の処理を行うが、第3符号C₃を得る処理を、 $i=3$ ($i=2, 3, \dots, n$ 、この例では4まで)を例として説明する。第 $i-1$ (=第2)符号C _{$i-1$} (=C₂)が第 $i-1$ (=第2)復号器25₂で復号されて標本化周波数2f _{$i-1$} (=2f₂)の第 $i-1$ (=第2)復号信号を得、この第 $i-1$ (=第2)復号信号と第 $i-2$ (=第1)サンプルレート変換器26 _{$i-2$} (=26₁)よりの第 $i-2$ (=第1)変換復号信号との和が加算器60 _{$i-1$} (=60₂)でとられ、その和信号は第 $i-1$ (=第2)サンプルレート変換器26 _{$i-1$} (=26₂)で標本化周波数2f _{i} (=2f₃)、周波数がf _{$i-1$} (=f₂)以下の第 $i-1$ (=第2)変換復号信号に変換される。一方、第 i (=第3)帯域選択手段としてのサンプルレート変換器22 _{i} (=22₃)により入力信号11から、周波数がf _{i} (=f₃)、標本化周波数が2f _{i} (=2f₃)の第 i (=第3)帯域信号23 _{i} (=23₃)が取出され、その第 i (=第3)帯域信号23 _{i} (=23₃)は第 $i-1$ (=第2)サンプルレート変換器26 _{$i-1$} (=26₂)よりの変換復号信号が第 i (=第3)差回路28 _{i} (28₃)で減算され、その第 i (=第3)減算信号29 _{i} が第 i (=第3)符号器24 _{i} (=24₃)で符号化され、第 i (=第3)符号C _{i} (=C₃)を出力する。なお、第 $i-1$ (=第2)復号器25 _{$i-1$} (=25₂)と、加算器60 _{$i-1$} (=60₂)と第 $i-1$ (=第2)サンプルレート変換器26 _{$i-1$} (=26₂)は第 $i-1$ (=第2)復号化手段40 _{$i-1$} (=40₂)を構成する。ただ第1復号化手段40₁は第 $i-2$ 層が存在せず加算器60₀は省略される。また最上位層、この例では第 i (=第4)帯域信号23₄は周波数f₄以下の信号であるため第 i 帯域選択手段としてのサンプルレート変換器22₄は省略される。

【0027】このようにしてこの発明は入力信号帯域をn区間に分割して符号化する場合に適用できる。第1～第n (=第4)符号C₁～C_n (=C₄)は多重化回路31でフレームごとに多重化されて符号化符号Cとして出力される。この場合多重化回路31は第1又は第1～第 i 符号の何れでも選択して出力することができるようにされる。第1～第n (=第4)符号器24₁～24_n (=24₄)は符号器24 _{i} の i が大となる程圧縮率が小さくなる、という使い方をすれば広帯域、高品質の符号化をする。これを満たせばその符号化方法は、例えば全てを変換符号化としてもよい。

【0028】第1～第4符号器24₁～24₄において聴覚重み付け符号化を行う場合はサンプルレート変換器22₁, 22₂, 22₃よりの各周波数がf₁, f₂, f₃以下の信号が聴覚重み付け係数演算部72₁, 72

2, 72₃へそれぞれ供給され、それぞれそのスペクトル包絡に基づく聴覚重み付け係数が演算され、また入力信号が聴覚重み付け係数演算部72₄に入力されて同様に聴覚重み付け係数が演算され、これら聴覚重み付け係数演算部72₁～72₄でそれぞれ演算された聴覚重み付け係数が第1～第4符号器24₁～24₄へ供給され、前述したように聴覚重み付け符号化が行われる。

【0029】この発明の符号化方法をn階層分割符号化への適用例としてn=4の場合を図10に示す。この例も原音入力信号11の最高周波数がf_n=f₄でその標本化周波数が2f₄の場合で、第1サンプルレート変換器(第1帯域選択手段)22₁で標本化周波数が2f₁(但しf₁<f₂<f₃<f₄)の入力信号23₁に変換され、つまり周波数f₁以下の第1帯域信号23₁が選出され、その第1帯域信号23₁は第1符号器24₁で符号化され、第1符号C₁として出力されると共にその第1符号C₁は第1復号器25₁で標本化周波数2f₁の信号に復号され、その復号信号12₁と第1帯域信号23₁との差が第1差回路65₁でとられ、その差信号(第1誤差信号)13₁は第1サンプルレート変換器26₁で標本化周波数が2f₂の第1変換誤差信号に変換される。

【0030】一方入力信号11から第2帯域選択手段66₂で周波数帯域がf₁～f₂、標本化周波数が2f₂の第2帯域信号23₂が取出される。例えば入力信号11がサンプルレート変換器22₂で標本化周波数2f₂の信号に変換され、その信号が遮断周波数f₁の高域通過フィルタ67₂に通されて第2帯域信号23₂が得られる。この第2帯域信号23₂は第1サンプルレート変換器26₁よりの第1変換誤差信号と第2加算器68₂で加算され、その第2加算信号29₂が第2符号器24₂で符号化され、第2符号C₂が出力される。

【0031】以下同様の処理を行うが、第3符号C₃を得る処理を、 $i=3$ ($i=2, 3, \dots, n$ 、この例では4まで)を例として説明する。第 $i-1$ (=第2)符号C _{$i-1$} (=C₂)が第 $i-1$ (=第2)復号器25₂で復号されて標本化周波数2f _{$i-1$} (=2f₂)の第 $i-1$ (=第2)復号信号を得、この第 $i-1$ (=第2)復号信号と第 $i-1$ (=第2)加算器68 _{$i-1$} (=68₂)より第 $i-1$ (=第2)加算信号29 _{$i-1$} (=29₂)との差が差回路65 _{$i-1$} (=65₂)でとられ、その第 $i-1$ (=第2)誤差信号13₂は第 $i-1$ (=第2)サンプルレート変換器26 _{$i-1$} (=26₂)で標本化周波数2f _{i} (=2f₃)の第 $i-1$ (=第2)変換誤差信号に変換される。一方、第 i (=第3)帯域選択手段66 _{i} (=66₃)により入力信号11から、帯域がf _{$i-1$} ～f _{i} (=f₂～f₃)、標本化周波数が2f _{i} (=f₃)の第 i (=第3)帯域信号23 _{i} (=23₃)が取出され、その第 i (=第3)帯域信号23 _{i} (=23₃)は第 $i-1$ (=第2)変換誤差信号と第 i

(=第3)加算器68_i (=68₃)で加算され、その第i (=第3)加算信号29₃が第i (=第3)符号器24_i (=24₃)で符号化され、第i (=第3)符号C_i (=C₃)を出力する。

【0032】このようにしてこの発明は入力信号帯域をn区間に分割して符号化する場合に適用できる。最上位層、つまり周波数 $f_{n-1} \sim f_n$ ($=f_3 \sim f_4$)の帯域を選出する第n (=第4)帯域選択手段66_n (=66₄)は単なる遮断周波数が f_{n-1} ($=f_3$)の高域通過フィルタ67_n (=67₄)でよい。第1~第n (=第4)符号C₁~C_n (=C₄)は多重化回路31でフレームごとに多重化されて符号化符号Cとして出力される。この場合多重化回路31は第1又は第1~第i符号の何れでも選択して出力することができるようにされる。

【0033】第1~第n (=第4)符号器24₁~24_n (=24₄)は符号器24_iのiが大となる程圧縮率が小さくなる、という使い方を行えば広帯域、高品質の符号化をする。これを満せばその符号化方法は、例えば全てを変換符号化としてもよい。第1~第4符号器24₁~24₄において聴覚重み付け符号化を行う場合はサンプルレート変換器71₁, 71₂, 71₃により入力信号がそれぞれ標本化周波数が2f₁, 2f₂, 2f₃の信号により変換されることにより、入力信号11からそれぞれ周波数がf₁, f₂, f₃以下の信号が取出されて聴覚重み付け係数演算部72₁, 72₂, 72₃へそれぞれ供給され、それぞれそのスペクトル包絡に基づく聴覚重み付け係数が演算され、また入力信号が聴覚重み付け係数演算部72₄に入力されて同様に聴覚重み付け係数が演算され、これら聴覚重み付け係数演算部72₁~72₄でそれぞれ演算された聴覚重み付け係数が第1~第4符号器24₁~24₄へ供給され、前述したように聴覚重み付け符号化が行われる。

【0034】この発明による復号化方法の一般的な方法を適用した復号化器の例として、n=4、つまり入力符号が第1~第4符号C₁~C₄が入力される場合を図11に図8と対応する部分に同一符号を付けて示す。符号分離手段56で入力符号Cは第1~第4符号C₁~C₄に分離されて、それぞれ第1~第4復号器57₁~57₄へ供給される。第1復号器57₁の第1復号信号58₁は第1復号化出力63₁として出力されると共にサンプルレート変換器59₁で標本化周波数がそれぞれ2f₂、第1変換復号信号61₁に変換され、その第1変換復号信号61₁は第2復号器57₂より第2復号信号58₂に第2加算器62₂で加算されて第2復号化出力63₂として出力されると共に第2サンプルレート変換器59₂で標本化周波数が2f₃の変換復号信号に変換される。一般には第i-1 (i=2, 3, ..., n、例えばi=3)加算器62_{i-1} (=62₂)よりの第i-1

(=第3)復号化出力63_{i-1} (=63₂)が第i-1 (第2)サンプルレート変換器59_{i-1} (=59₂)で標本化周波数が2f_i (=2f₃)の第i-1 (=第2)変換復号信号61_{i-1} (=61₂)に変換され、その第i-1 (=第2)変換復号信号61_{i-1} (=61₂)と第i (=第3)復号器57_i (=57₃)からの第i (=第3)復号信号58_i (=58₃)とが第i (=第3)加算器62_i (=62₃)で加算されて第i (=第3)復号化出力63_i (=63₃)を得、これが出力される。

【0035】

【発明の効果】以上説明したように、この発明によれば、階層符号化方法において下位層の量子化誤差を上位層で符号化しているため、CELP符号化方法と変換符号化方法などの、圧縮方法の異なる符号化方法によって階層を構成しても、上位層までの復号信号において符号化品質を低下させない、という効果がある。また、下位層の量子化誤差を上位層で符号化する、という操作を繰り返すことにより、複数階層化において量子化誤差を階層数に応じて減少させることが可能となる。更に、このような符号化方法によって、どの階層で復号しても聴感上の復号品質が最適となり、スケーラブルな階層符号化を実現できる。

【図面の簡単な説明】

【図1】サブバンド符号化方法を3つの周波数帯域に分割する方法によって実現した場合の原音(A)と符号化再生音(B)、および量子化誤差(C)の例を示す図。

【図2】スケーラブルな階層構造を持つ階層符号化方法の特徴を説明するための図。

【図3】サブバンド符号化方法によって階層符号化を実現した場合の原音、復号信号、量子化誤差の様子を示す図。

【図4】Aはこの発明による符号化方法を2階層符号化法に適用した場合の符号化器の例を示すブロック図、Bは多重化された符号の例を示す図である。

【図5】A~Dは図4Aの符号化動作における原音、復号信号、上位層符号化入力、上位層聴覚重み付けの基準の各例を示す図、E、Fは上位層の復号信号、上位層までの復号の量子化誤差の例を示す図である。

【図6】CELP符号化器の概略を示すブロック図。

【図7】変換符号化器の概略を示すブロック図。

【図8】この発明の復号化方法を2階層符号化の復号法に適用した復号器の例を示すブロック図。

【図9】この発明の符号化方法を4階層符号化方法として実現した場合の符号器の例を示すブロック図。

【図10】この発明による4階層符号化方法を実現する符号器の他の例を示すブロック図。

【図11】この発明の復号化方法を4階層符号化方法として実現した場合の復号器の例を示すブロック図。

【図1】

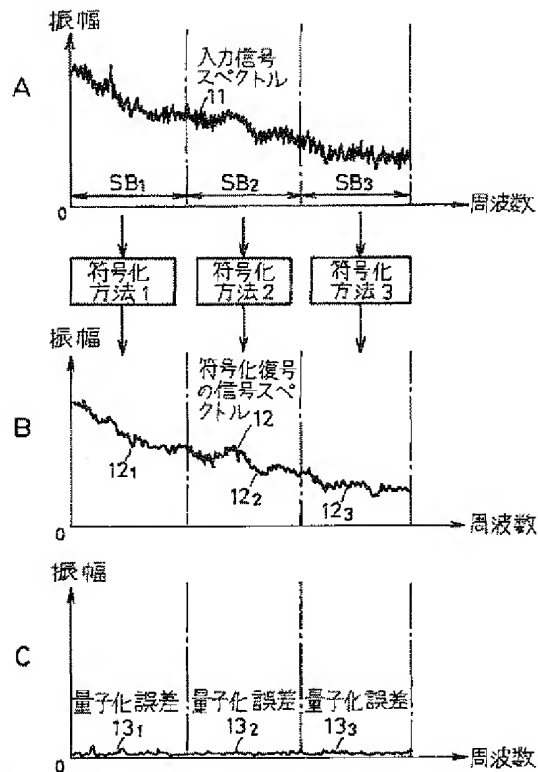


図1

【図7】

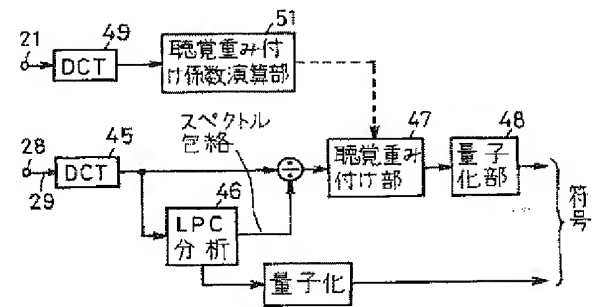


図7

【図2】

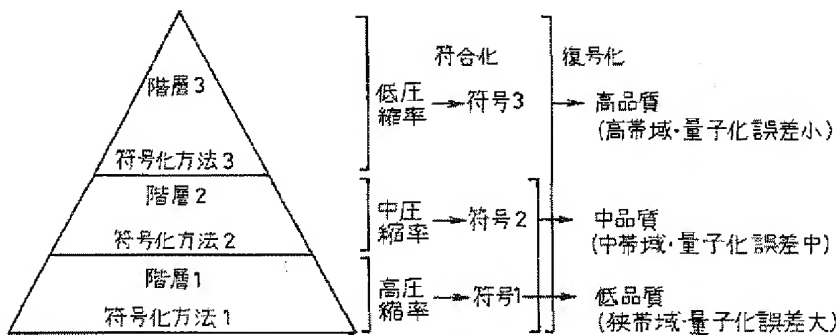


図2

【図3】

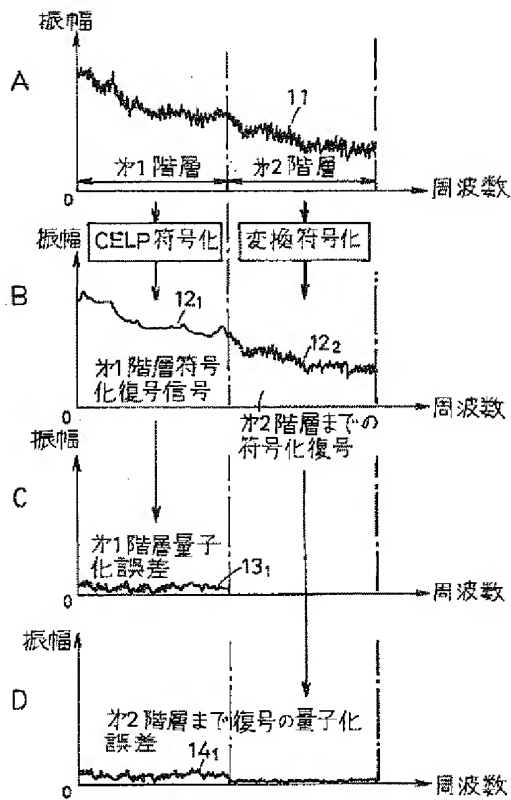


図3

【図6】

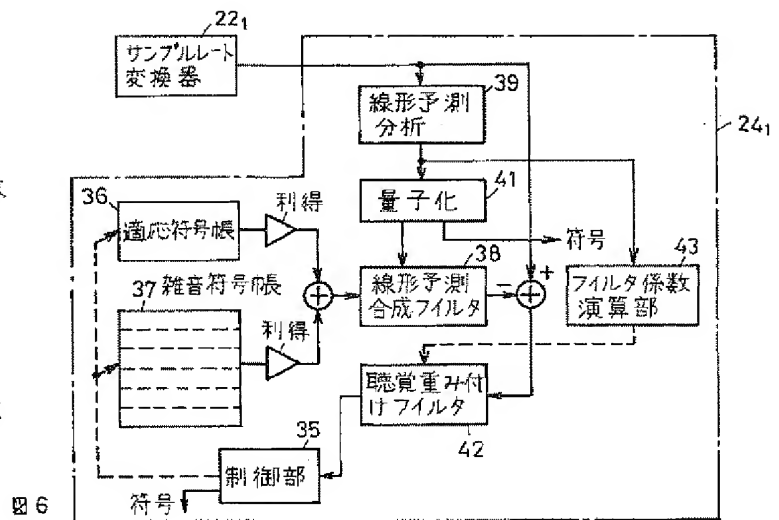


図6

【図4】

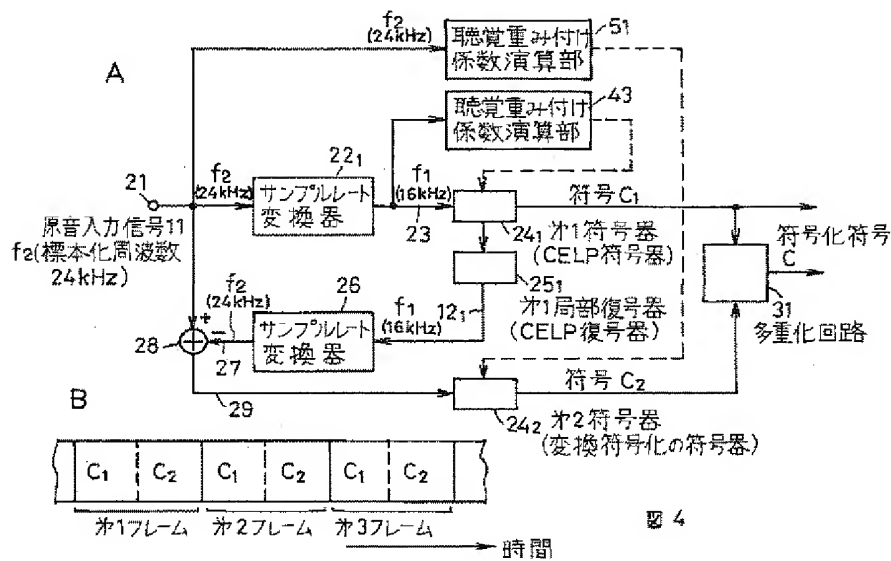


図4

【図5】

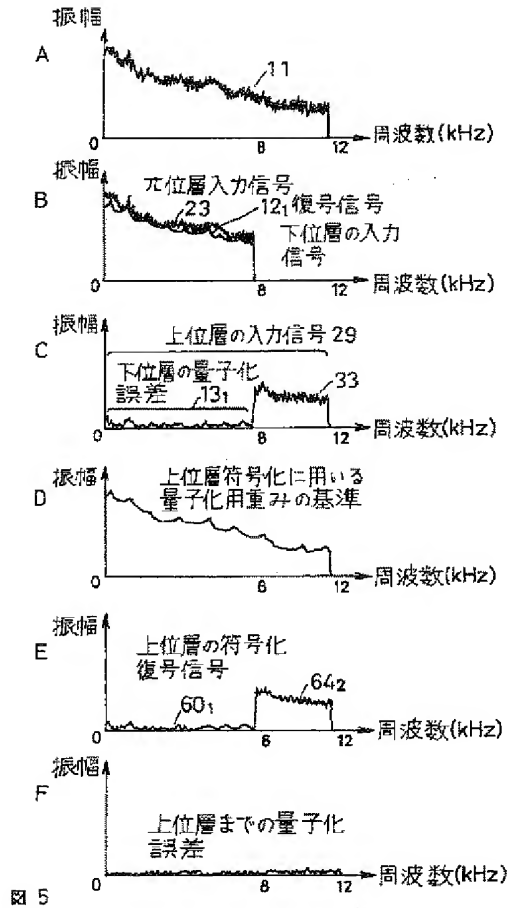


図 5

【図8】

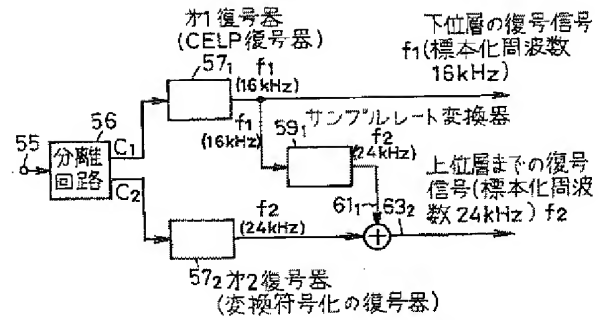


図 8

【図9】

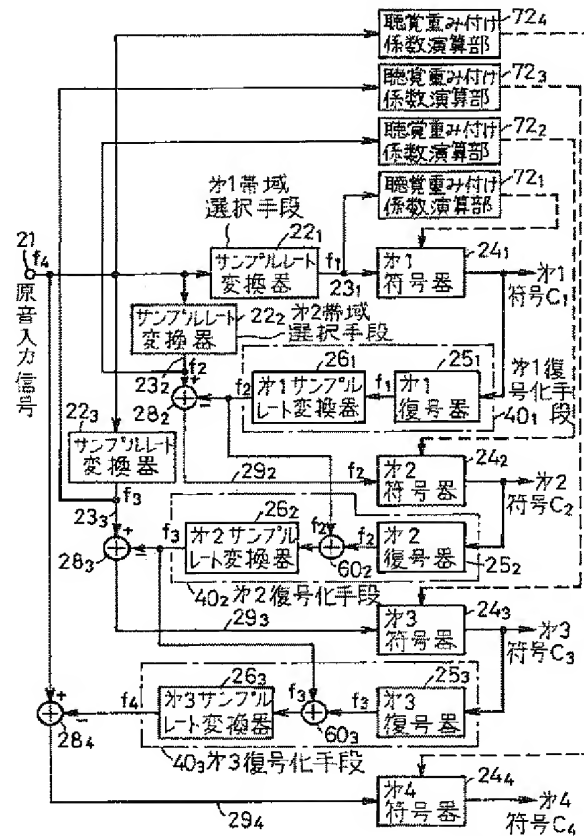


図 9

【図10】

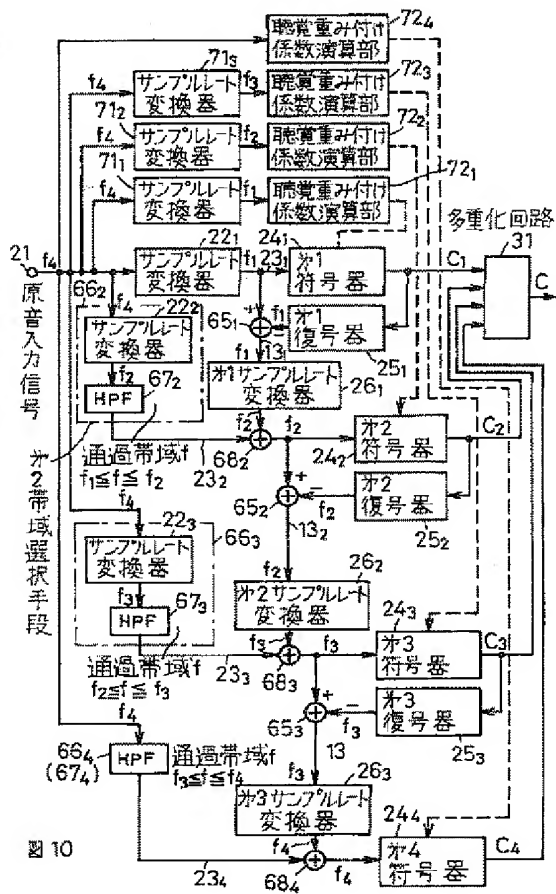


図 10

【図11】

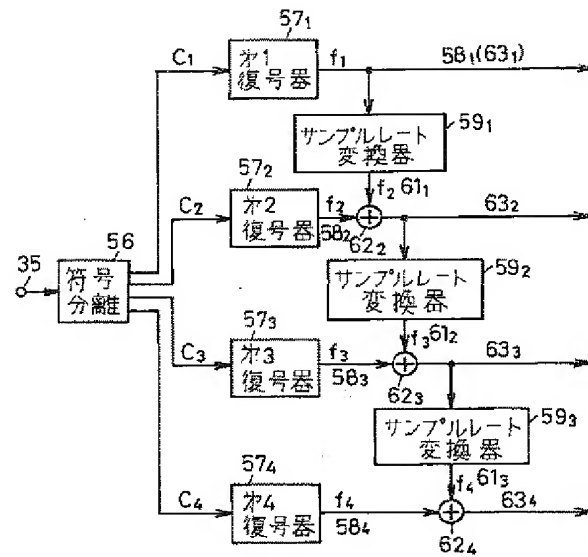


図 11

[0002]

[Description of the Related Art]

A subband coding method is available as a method of coding an acoustic signal by dividing its band in a frequency domain. The subband coding method divides an input signal into a plurality of frequency bands using QMF (Quadrature Mirror Filter) and codes the bands independently of each other while assigning appropriate bits to the respective bands.

[0003] A variety of methods for coding acoustic signals such as

musical sound and speech are currently available depending on the purpose of use, decoding quality, coding speed or the like, but it is common practice that one acoustic signal is coded not using a plurality of coding methods but using only one coding method.

However, for example, as shown in FIG.1A, acoustic signal 11 is divided into three subbands SB_1 , SB_2 and SB_3 from the low-frequency side on the frequency axis, thereby stratified, and

as shown in FIG.2, subband SB_1 which is a lower layer (layer 1) is coded using a coding method of low coding quality, that is, having a narrow frequency band of decoded and reproduced sound and

having large quantization errors, for example, code excited linear predictive coding method: CELP at a high compression rate, and by contrast, subband SB_3 which is a higher layer (layer 3) is coded using a coding method of high coding quality, that is, having a wide frequency band of decoded and reproduced sound and having

small quantization errors (e.g., transform coding method such as discrete cosine transform coding method) at a low compression rate and subband SB_2 which is an intermediate layer (layer 2) is coded using a coding method intermediated between the lower layer coding method and the higher layer coding method, and

different coding methods may also be adopted at the request of a user such as coding and transmitting only layer 1 or coding and transmitting layers 1 and 2 or coding and transmitting all layers.

[0004] Alternatively, as described above, various musical sound or speech signals coded in three layers may be provided, for example, as a database and the user may access the database, receive desired musical sound signals, decode only the code of layer 1 according to the user's decoder, obtain reproduced sound

of low quality in a narrow band with large quantization errors or decode codes of layers 1 and 2 or decode all codes of layers 1, 2 and 3 to obtain reproduced sound of high quality in a wideband and with small quantization errors.

5 [0005] Alternatively, for example, if an acoustic signal in a wideband where speech is predominant is divided into two layers and coded and only the lower layer code is decoded, the acoustic signal having speech-like nature may be decoded clearly or if the codes of both the lower layer and the higher layer are decoded,
10 signals including the acoustic signal having non-speech-like nature may be decoded. Furthermore, in these cases, a decoded signal of high quality can also be obtained by receiving only the lower layer code, shortening the time using a transmission path or using a transmission path of small transmission capacity, and
15 decoding the code in real time or receiving also the higher layer code for a long time, storing the code once and then reproducing and decoding the code.

[0006] Alternatively, in these cases, it is possible to store all the codes of the lower and higher layers once, then decode only the
20 lower layer code using a small, economic decoder with a small delay time in real time or when high quality sound is required to be reproduced, it is possible to decode the sound including the higher layer code using a large decoder with a large delay time by taking a longer time and then reproduce the sound at one time later.

25 [0007] As described above, the coding method that provides selectivity for decoding quality or a coding compression rate is called a "scalable layered coding method." As such a scalable layered coding method, the subband coding method shown in FIG.1A may be used. That is, the frequency band of subband SB_1
30 is coded using coding method 1 and bands SB_2 and SB_3 are likewise coded using coding methods 2 and 3 which are independent of each other. As shown in FIG.1B, at the time of decoding, when, for example, wideband decoded sound is not necessary, only the code of subband SB_1 is decoded using a decoder of coding method 1,
35 decoded signal 12_1 of the sound of only the band of subband SB_1 is obtained, and on the other hand when wideband decoded sound is necessary, the codes of subbands SB_1 , SB_2 and SB_3 are decoded

using a decoder corresponding to coding methods 1, 2 and 3, decoded signals 12_1 , 12_2 and 12_3 are thereby obtained and signal 12 synthesizing these signals is outputted.

[0008]

5 [Problems to be Solved by the Invention]

However, in layered coding using such a subband coding method, a quantization error produced in each band (that is, each layer), that is, an error between an input signal of a coder and an output signal of a local decoder thereof, that is, a decoded signal
10 unaffected by transmission paths or the like may be stored in each band SB_1 , SB_2 or SB_3 , as quantization error 13_1 , 13_2 or 13_3 respectively as shown in FIG.1C and distortion or noise may be generated in decoded signal 12 of the entire frequency band independently of each other in each band. Therefore, even when
15 the entire band is decoded (that is, up to the higher layer is decoded), large quantization error 13_1 of the lower layer is also generated as is, and therefore high quality signals cannot be obtained. It is not possible to reduce quantization noise in order to obtain a wideband decoded signal of high quality unless each
20 coding compression rate of subband SB_1 , SB_2 or SB_3 is reduced. Therefore, such a layered coding method cannot realize scalable coding.

[0009] The reason that scalable coding cannot be realized using the conventional subband coding method will be described more
25 specifically with reference to FIG.3. That is, a band of original acoustic signal 11 is divided into two portions, and a first layer (low-frequency region) is coded based on a CELP scheme and a second layer (high-frequency region) is coded using a transform coding method. In the first layer, since CELP coding of high
30 speech compression efficiency is executed, quantization error signal 13_1 of local decoded signal 12 (FIG.3B) is relatively large as shown in FIG.3C. On the other hand, in the second layer, since transform coding which is capable of coding various waveforms is executed, local decoded signal 12_2 is close to original sound
35 signal 11 as shown in FIG.3B and quantization error signal 13_2 is small as shown in FIG.3C. However, even if the code of the first layer and the code of the second layer are decoded and wideband

decoded signals are obtained, low-frequency portion 14₁ of the quantization error of the decoded signal is not different from quantization error 13₁ of the first layer as shown in FIG.3D. That is, the decoding quality up to the second layer depends on the coding performance of the CELP coding method in a low-frequency band. Therefore, in order for the subband coding method to perform layered coding and realize high coding quality, it is necessary to perform coding on all layers using a high quality coding method at a small compression rate or with a large amount of calculation.

[0010] It is therefore an object of the present invention to provide a scalable coding method and a decoding method thereof capable of performing coding in a lower layer at a high compression rate and with low decoding quality and obtaining a high quality decoded signal up to a higher layer unaffected by low decoding quality of the lower layer.

[0011]

[Means for Solving the Problem]

The invention according to claim 1 is a coding method of performing coding by dividing an acoustic input signal such as musical sound and speech having a highest frequency of f_n into n segments of frequency f_1, f_2, \dots, f_{n-1} ($f_1 < f_2 < \dots < f_{n-1} < f_n$) (n is an integer of 2 or greater), including the steps of selecting a first band signal having a frequency of f_1 or less from the input signal, coding the first band signal and outputting a first code using a first coding method, obtaining an $(i-1)$ th decoded signal having a frequency of f_{i-1} or less from each of $(i-1)$ th or lower numbered codes ($i=2, 3, \dots, n$), selecting an i th band signal having a frequency f_i or less from the above described input signal, subtracting the $(i-1)$ th decoded signal from the i th band signal and obtaining an i th difference signal, and coding the i th difference signal using an i th coding method and outputting an i th code.

[0012] The $(i-1)$ th decoded signal is obtained, for example, by adding up a signal obtained by decoding the $(i-1)$ th code and the $(i-2)$ th decoded signal and converting the addition signal to a signal having a sampling frequency of $2f_i$. The coding method of the invention according to claim 3 is a coding method of

performing coding by dividing an acoustic input signal such as musical sound and speech having a highest frequency of f_n into n segments of frequency f_1, f_2, \dots, f_{n-1} ($f_1 < f_2 < \dots < f_{n-1} < f_n$) (n is an integer of 2 or greater), including the steps of obtaining a first band acoustic signal having a sampling frequency of $2f_1$ from the above described acoustic input signal, coding the first band signal using a first coding method and outputting a first code, obtaining a coding error of the first code as an $(i-1)$ th error signal ($i=2, 3, \dots, n$), converting the $(i-1)$ th error signal to an $(i-1)$ th converted error signal having a sampling frequency of $2f_i$, obtaining an i th band signal having a frequency band of f_{i-1} to f_i , and a sampling frequency of $2f_i$ from the above described acoustic input signal, adding up the above described $(i-1)$ th converted error signal and the above described i th band signal and obtaining an i th addition signal, coding the i th addition signal using an i th coding method and outputting an i th code.

[0013] The invention according to claim 4 is one of the inventions according to claims 1 to 3, wherein the above described first coding method is a code excited linear predictive coding method and the above described n th coding method is a transform coding method. The invention according to claim 5 is one of the inventions according to claims 1 to 4, wherein psychological perceptual weighted quantization is performed in a coding step of the above described i th code using a spectrum envelope of components of frequency f_i or less in substantially the whole band of the above described acoustic input signal.

[0014] The decoding method of the invention according to claim 6 separates an input code into first to n th codes (n is an integer of 2 or greater), decodes the above described first code, outputs a first decoded signal having a sampling frequency of $2f_1$, converts the above described $(i-1)$ th decoded signal ($i=2, 3, \dots, n$) to an $(i-1)$ th converted decoded signal having a sampling frequency of $2f_i$, decodes the above described i th code, obtains an i th decoded signal having a sampling frequency of $2f_i$, adds up the i th decoded signal and the above described $(i-1)$ th converted decoded signal and outputs an i th addition signal.

[0015]

[Embodiment]

FIG.4A shows an example of a coder to which an embodiment of the coding method of the invention according to claim 1 is applied. This example describes a case where an original sound signal is divided into two frequency bands and coded, that is, two-layered coding is performed. Original sound input signal 11 from input terminal 21 is a digital signal having a sampling frequency of 24 kHz, that is, having highest frequency f_2 of 12 kHz and this input signal is converted to a signal having a sampling frequency of 16 kHz by sample rate converter 22, as a first band selection section and first band signal 23 is thereby extracted. This sample rate conversion is so-called down sampling and executed, for example, after removing samples at intervals corresponding to a conversion sampling frequency ratio by causing the signal to pass through a digital low pass filter. First band signal 23 having frequency f_1 of 8 kHz or less is extracted from sample rate converter 22, and first band signal 23 is coded by first coder 24, according to a first coding method. In this example, first coder 24, performs coding using a CELP (code excited linear prediction coding) coding method. First code C_1 which is this coding result is outputted.

[0016] In the present embodiment, local decoder 25, performs decoding to obtain first decoded signal 12, having a frequency of f_1 or less and first sample rate converter 26, converts decoded signal 12, to converted decoded signal 27 having a sampling frequency of 24 kHz. Sample rate converter 26, is intended to perform so-called up sampling, and, for example, may add 0 samples according to a conversion frequency ratio and then cause the signal to pass through a digital low pass filter. Difference circuit 28 subtracts converted decoded signal 27 from input signal 11 and second coder 24₂ codes difference signal 29 using a second coding method. In the present embodiment, second coder 24₂ performs coding through transform coding such as modified discrete cosine transform. Second code C_2 of this coding result is outputted. First code C_1 and second code C_2 are multiplexed on a time-division basis for each coding frame by multiplexing circuit

31 as shown, for example, in FIG.4b and outputted as coded code C. Only first code C_1 may be outputted at the request of the user.

[0017] A frequency spectrum of original sound input signal 11 having a sampling frequency of 24 kHz is shown, for example, in FIG.5A, signals of 8 kHz or less of this signal 11 are inputted to first coder 24₁ of the lower layer as signal 23 (FIG.5B) having a sampling frequency of 16 kHz and coded with high compression efficiency. In decoded signal 12₁ decoded from coded code C_1 by local decoder 25₁, quantization error 13₁ occurs to no small extent as shown in FIG.5C with respect to lower layer input signal 23 as shown in FIG.5B. From difference circuit 28, signal 29 made up of error signal 13₁ and high-frequency signal 33 of 8 kHz or higher of original sound input signal 11 is inputted to second coder 24₂ of the higher layer and coded using a transform coding method capable of coding input signals of all nature with high quality.

[0018] Thus, according to the present embodiment, coded code C_1 in the lower layer does not code the original sound so faithfully, but coded code C_1 in the higher layer codes the original sound including a quantization error of the lower layer, and therefore when decoding is performed up to the higher layer, the lower layer can also be decoded and reproduced with high fidelity as will be made clear later. That is, coding in the lower layer can be performed with high compression efficiency and a high quality decoded signal can be obtained when the higher layer is also decoded.

[0019] Since the embodiment in particular uses a CELP scheme for coding of the lower layer, when the coding target is speech, even if only first code C_1 in the lower layer is decoded, relatively high quality is obtained, the amount of calculation is small and real-time processing is performed easily. When first and second codes C_1 and C_2 are decoded and even if the coding target is musical sound, a decoded signal of high quality is obtained over a wideband through decoding of the transform code in the higher layer and compensation of coding errors of the CELP code in the lower layer.

[0020] When coding is performed, it is often the case that coding is performed by assigning psychological perceptual weights in

consideration of human psychological perception, for example, masking characteristics with a large level spectrum, and efficient coding with perceptually suppressed quantization errors is thereby performed. For example, in the CELP coding method of coder 24₁, as shown in FIG.6, a vector of a cycle (pitch) specified by control section 35 is extracted from adaptive codebook 36, furthermore a noise vector is extracted from specified noise codebook 37, these vectors are assigned gains respectively, synthesized and inputted to linear predictive synthesis filter 38 as excitation vectors. On the other hand, the input signal from sample rate converter 22₁ in FIG.4A is subjected to a linear predictive analysis by linear predictive analysis section 39 in a coding frame cycle, the linear prediction coefficient thereof is quantized by quantization section 41 and a filter coefficient of synthesis filter 38 is set according to the quantization linear prediction coefficient. Furthermore, perceptual weighted coefficient calculation section 43 calculates a filter coefficient for psychological perceptual weighting based on the spectrum envelope calculated from the linear prediction coefficient and sets the filter coefficient in perceptual weighted filter 42. A synthesized signal from synthesis filter 38 is subtracted from the input signal from sample rate converter 22₁, the difference signal thereof is passed to perceptual weighted filter 42, and control section 35 performs selections on adaptive codebook 36 and noise codebook 37 so that the output energy thereof becomes a minimum.

[0021] According to the transform coding method of transform coder 24₂, for example, the output of difference circuit 28 is subjected to orthogonal cosine transform by discrete cosine transformer 45 as shown in FIG.7, transformed to a coefficient in the frequency domain, the spectrum envelope component thereof is subjected to a linear predictive analysis at linear predictive analysis section 46, a spectrum envelope is thereby obtained, the output coefficient of transformer 45 is divided by the spectrum envelope and normalized, the averaged coefficient is assigned a perceptual weight by perceptual weighting section 47, and further subjected to, for example, vector quantization by quantization section 48. To obtain a perceptual weighted coefficient, in the

present embodiment, original sound input signal 11 from input terminal 21 is subjected to orthogonal cosine transform by discrete cosine transformer 49, transformed into the frequency domain, a perceptual weighted coefficient is calculated by coefficient calculation section 51 based on the spectrum envelope of the transform coefficient, given to perceptual weighting section 47 and the corresponding component of the normalization coefficient is multiplied by the perceptual weighted coefficient.

[0022] That is, second coder 24₂ of the higher layer codes signal 29 of the spectrum shown in FIG.5C, but instead of performing perceptual weighting based on the spectrum envelope of this signal 29, second coder 24₂ calculates a spectrum envelope of original sound input signal 11 (FIG.5D), and performs perceptual weighted coding based on this. Next, the present embodiment of the decoding method of the present invention will be described with reference to FIG.8. The present embodiment is a case where the present invention is applied to decoding on the coded code using the coding method shown in FIG.4. An input code inputted from input terminal 55 is demultiplexed into first code C₁ and second code C₂ by demultiplexing circuit 56, first code C₁ is decoded into first decoded signal 58₁ of highest signal frequency f₁ (sampling frequency 16 kHz) by first decoder 57₁ using a CELP decoding method in this example and outputted as lower layer (low-frequency) decoded output 63₁.

[0023] This first decoded output 58₁ is converted to converted decoded signal 61₁ of highest signal frequency f₂ (sampling frequency is 24 kHz) by sample rate converter 59. On the other hand, second code C₂ from demultiplexing circuit 56 is subjected to transform code decoding in this example by second decoder 57₂, second decoded signal 58₂ of highest signal frequency f₂ (sampling frequency is 24 kHz) is obtained, and this second decoded signal 58₂ is added to first converted decoded signal 61₁ by adder 62₂ and outputted as higher layer (whole band) decoded output 63₂.

[0024] That is, decoded signal 12₁ in FIG.5B is obtained as lower layer decoded output 63₁ in an ideal case. On the other hand, decoded signal 58₂ of second decoder 57₂ is decoded signal 60₁ of quantization error signal 13₁ of the lower layer (low-frequency)

and decoded signal 64_2 of high-frequency signal 33 as shown in FIG.5E in an ideal case. Therefore, decoded signal 60_1 corresponding to quantization error 13_1 for low-frequency decoded signal 58_1 is added to decoded output 63_2 from adder 62_2 and
5 quantization errors are noticeably reduced and demonstrates a high degree of fidelity with respect to high-frequency decoded signal 64_2 , and therefore, decoded output 63_2 up to the higher layer obtained from adder 62 is significantly close to original sound input signal 11 and the quantization error signal thereof is
10 significantly small over the whole band as shown, for example, in FIG.5F.

[0025] Next, a case with $n=4$ will be described as an example where the coding method of the present invention is applied to n -layer (n -band) divided coding with reference to FIG.9. In FIG.9,
15 parts corresponding to those in FIG.4A are assigned the same reference numerals. In this example, original sound input signal 11 has a highest frequency of $f_n=f_4$ and a sampling frequency of $2f_4$, first sample rate converter (first band selection section) 22_1 converts original sound input signal 11 to input signal 23_1 having a sampling frequency of $2f_1$ (where $f_1 < f_2 < f_3 < f_4$), that is, first band signal 23_1 of frequency f_1 or less is selected, first band signal 23_1 is coded by first coder 24_1 , outputted as first code C_1 and first code C_1 is decoded into a signal of sampling frequency $2f_1$ by first decoder 25_1 and decoded signal 12_1 is converted to a first
20 converted decoded signal having a sampling frequency of $2f_2$ by first sample rate converter 26_1 . On the other hand, input signal 11 is converted to a signal having a sampling frequency of $2f_2$ by sample rate converter 22_2 as the second band selection section, and second band signal 23_2 of frequency f_2 or less is extracted.
25 Second difference circuit 28_2 subtracts the first converted decoded signal from first sample rate converter 26_1 from second band signal 23_2 and second coder 24_2 codes second difference signal 29_2 and outputs second code C_2 .

[0026] Hereinafter, similar processing will be performed, and
35 the processing of obtaining third code C_3 in a case with $i=3$ ($i=2, 3, \dots, n$, up to 4 in this example) will be described as an example. $(i-1)$ th (=second) code C_{i-1} (= C_2) is decoded by $(i-1)$ th (=second)

decoder 25_2 , an $(i-1)$ th (=second) decoded signal of sampling frequency $2f_{i-1}$ ($=2f_2$) is obtained, adder 60_{i-1} ($=60_2$) calculates the sum of the $(i-1)$ th (=second) decoded signal and $(i-2)$ th (=first) converted decoded signal from $(i-2)$ th (=first) sample rate converter 26_{i-2} ($=26_1$), the sum signal is converted to an $(i-1)$ th (=second) converted decoded signal of sampling frequency $2f_i$ ($=2f_3$) and frequency f_{i-1} ($=f_2$) or less by $(i-1)$ th (=second) sample rate converter 26_{i-1} ($=26_2$). On the other hand, sample rate converter 22_i ($=22_3$) as an i th (=third) band selection section extracts i th (=third) band signal 23_i ($=23_3$) having a frequency of f_i ($=f_3$) and a sampling frequency of $2f_i$ ($=2f_3$) from input signal 11, i th (=third) difference circuit 28_i ($=28_3$) subtracts the converted decoded signal from $(i-1)$ th (=second) sample rate converter 26_{i-1} ($=26_2$) from i th (=third) band signal 23_i ($=23_3$) and i th (=third) coder 24_i ($=24_3$) codes i th (=third) subtracted signal 29_3 and outputs i th (=third) code C_i ($=C_3$). $(i-1)$ th (=second) decoder 25_{i-1} ($=25_2$), adder 60_{i-1} ($=60_2$) and $(i-1)$ th (=second) sample rate converter 26_{i-1} ($=26_2$) constitute $(i-1)$ th (=second) decoding section 40_{i-1} ($=40_2$). However, first decoding section 40_1 has no $(i-2)$ th layer, and adder 60_0 is omitted. Furthermore, since the highest layer, i th (=fourth) band signal 23_4 in this example is a signal having frequency f_4 or less, sample rate converter 22_4 as the i th band selection section is omitted.

[0027] By this means, the present invention is applicable to a case where coding is performed by dividing an input signal band into n segments. First to n th (=fourth) codes C_1 to C_n ($=C_4$) are multiplexed by multiplexing circuit 31 for each frame and outputted as coded code C . In this case, multiplexing circuit 31 is enabled to select and output a first code or one of first to i th codes. When used in such a way that the greater i of coder 24_i , the smaller the compression rate becomes, first to n th (=fourth) coders 24_1 to 24_n ($=24_4$) perform wideband and high quality coding. In so far as this is satisfied, all the coding method may be realized by, for example, transform coding.

[0028] When first to fourth coders 24_1 to 24_4 perform perceptual weighted coding, a signal having each frequency of f_1 , f_2 , f_3 or less from sample rate converters 22_1 , 22_2 , and 22_3 is supplied to

perceptual weighted coefficient calculation section 72_1 , 72_2 , or 72_3 respectively, perceptual weighted coefficients based on a spectrum envelope are calculated respectively, and an input signal is inputted to perceptual weighted coefficient calculation section 72_4 , and perceptual weighted coefficients are likewise calculated, and the perceptual weighted coefficients calculated in perceptual weighted coefficient calculation sections 72_1 to 72_4 are supplied to first to fourth coders 24_1 to 24_4 and perceptual weighted coding is performed as described above.

[0029] FIG.10 shows a case with $n=4$ as an example of application of the coding method of the present invention to n -layer divided coding. Also in this example, when original sound input signal 11 has a highest frequency of $f_n = f_4$ and a sampling frequency of $2f_4$, first sample rate converter (first band selection section) 22_1 converts original sound input signal 11 to input signal 23_1 having a sampling frequency of $2f_1$ (where $f_1 < f_2 < f_3 < f_4$), that is, first band signal 23_1 of frequency f_1 or less is selected, first band signal 23_1 is coded by first coder 24_1 , outputted as first code C_1 and first code C_1 is decoded into a signal of sampling frequency $2f_1$ by first decoder 25_1 , and the difference between decoded signal 12_1 and first band signal 23_1 is calculated by first difference circuit 65_1 and difference signal (first error signal) 13_1 is converted to a first conversion error signal having a sampling frequency of $2f_2$ by first sample rate converter 26_1 .

[0030] On the other hand, second band selection section 66_2 extracts second band signal 23_2 having a frequency band of f_1 to f_2 , and a sampling frequency of $2f_2$ from input signal 11. For example, input signal 11 is converted to a signal having sampling frequency $2f_2$ by sample rate converter 22_2 , the signal is passed through high pass filter 67_2 having cutoff frequency f_1 and second band signal 23_2 is obtained. This second band signal 23_2 is added to the first conversion error signal from first sample rate converter 26_1 by second adder 68_2 , and second addition signal 29_2 is coded by second coder 24_2 , and second code C_2 is outputted.

[0031] Hereinafter, similar processing will be performed, and the processing of obtaining third code C_3 will be described taking a case with $i=3$ ($i=2, 3, \dots, n$, up to 4 in this example) as an

example. $(i-1)$ th (=second) code C_{i-1} ($=C_2$) is decoded by $(i-1)$ th (=second) decoder 25₂, an $(i-1)$ th (=second) decoded signal having sampling frequency $2f_{i-1}$ ($=2f_2$) is obtained, difference circuit 65_{i-1} ($=65_2$) calculates the difference between the $(i-1)$ th (=second) decoded signal and $(i-1)$ th (=second) addition signal 29_{i-1} ($=29_2$) from $(i-1)$ th (=second) adder 68_{i-1} ($=68_2$) and $(i-1)$ th (=second) error signal 13₂ is converted to $(i-1)$ th (=second) conversion error signal having sampling frequency $2f_i$ ($=2f_3$) by $(i-1)$ th (=second) sample rate converter 26_{i-1} ($=26_2$). On the other hand, i th (=third) band selection section 66_i ($=66_3$) extracts i th (=third) band signal 23_i ($=23_3$) having a band of f_{i-1} to f_i ($=f_2$ to f_3) and a sampling frequency of $2f_i$ ($=f_3$) from input signal 11, i th (=third) band signal 23_i ($=23_3$) is added to the $(i-1)$ th (=second) conversion error signal by i th (=third) adder 68_i ($=68_3$), i th (=third) addition signal 29₃ is coded by i th (=third) coder 24_i ($=24_3$) and i th (=third) code C_i ($=C_3$) is outputted.

[0032] By this means, the present invention is applicable to a case where coding is performed by dividing an input signal band into n segments. n th (=fourth) band selection section 66_n ($=66_4$) that selects the highest layer, that is, a band of frequency f_{n-1} to f_n ($=f_3$ to f_4) may be simply high pass filter 67_n ($=67_4$) having a cutoff frequency of f_{n-1} ($=f_3$). First to n th (=fourth) codes C_1 to C_n ($=C_4$) are multiplexed by multiplexing circuit 31 for each frame and outputted as coded code C . In this case, multiplexing circuit 31 is enabled to select and output the first code or one of first to i th codes.

[0033] When used in such a way that the greater i of coder 24_i, the smaller the compression rate becomes, first to n th (=fourth) coders 24₁ to 24_n ($=24_4$) perform wideband and high quality coding. In so far as this is satisfied, all the coding method may be realized by, for example, transform coding. When first to fourth coders 24₁ to 24₄ perform perceptual weighted coding, sample rate converters 71₁, 71₂ and 71₃ convert the input signals to signals having a sampling frequency of $2f_1$, $2f_2$ and $2f_3$, signals having a frequency of f_1 , f_2 , f_3 or less are thereby extracted from input signal 11, supplied to perceptual weighted coefficient calculation sections 72₁, 72₂ and 72₃ respectively, perceptual weighted

coefficients based on their respective spectrum envelopes are calculated, and an input signal is inputted to perceptual weighted coefficient calculation section 72₄, and perceptual weighted coefficients are likewise calculated, perceptual weighted coefficients calculated by perceptual weighted coefficient calculation sections 72₁ to 72₄ are supplied to first to fourth coders 24₁ to 24₄ and perceptual weighted coding is performed as described above.

[0034] FIG.11 shows a case with $n=4$, that is, a case where input codes of first to fourth codes C_1 to C_4 are inputted as an example of a decoder to which a general method of the decoding method according to the present invention is applied with parts corresponding to those in FIG.8 assigned the same reference numerals. Input code C is demultiplexed by code demultiplexing section 56 into first to fourth codes C_1 to C_4 and supplied to first to fourth decoders 57₁ to 57₄, respectively. First decoded signal 58₁ of first decoder 57₁ is outputted as first decoded output 63₁ and converted to a signal having a sampling frequency of $2f_2$ and first converted decoded signal 61₁ respectively by sample rate converter 59₁, and first converted decoded signal 61₁ is added to second decoded signal 58₂ from second decoder 57₂ by second adder 62₂, outputted as second decoded output 63₂ and converted to a converted decoded signal having a sampling frequency of $2f_3$ by second sample rate converter 59₂. Generally, $(i-1)$ th ($=$ third) decoded output 63 _{$i-1$} ($=63_2$) from $(i-1)$ th ($i=2, 3, \dots, n$, for example, $i=3$) adder 62 _{$i-1$} ($=62_2$) is converted to $(i-1)$ th ($=$ second) converted decoded signal 61 _{$i-1$} ($=61_2$) having a sampling frequency of $2f_i$ ($=2f_3$) by $(i-1)$ th (second) sample rate converter 59 _{$i-1$} ($=59_2$), $(i-1)$ th ($=$ second) converted decoded signal 61 _{$i-1$} ($=61_2$) and i th ($=$ third) decoded signal 58 _{i} ($=58_3$) from i th ($=$ third) decoder 57 _{i} ($=57_3$) are added up by i th ($=$ third) adder 62 _{i} ($=62_3$) and i th ($=$ third) decoded output 63 _{i} ($=63_3$) is thereby obtained and outputted.

FIG.1

A

AMPLITUDE

11 INPUT SIGNAL SPECTRUM

5 FREQUENCY

CODING METHOD 1

:

B

10 AMPLITUDE

12 CODED/DECODED SIGNAL SPECTRUM

FREQUENCY

C

15 AMPLITUDE

13, QUANTIZATION ERROR

:

FREQUENCY

20 FIG.2

LAYER 3

CODING METHOD 3

:

CODING

25 LOW COMPRESSION RATE

INTERMEDIATE COMPRESSION RATE

HIGH COMPRESSION RATE

CODE 3

:

30 DECODING

HIGH QUALITY

(HIGH-FREQUENCY BAND WITH SMALL QUANTIZATION
ERROR)

INTERMEDIATE QUALITY

35 (INTERMEDIATE BAND WITH INTERMEDIATE QUANTIZATION
ERROR)

LOW QUALITY

(NARROW BAND WITH LARGE QUANTIZATION ERROR)

FIG.3

A

5 AMPLITUDE
FIRST LAYER
SECOND LAYER
FREQUENCY

10 B

AMPLITUDE
CELP CODING
TRANSFORM CODING
FIRST LAYER CODED/DECODED SIGNAL
15 CODING/DECODING UP TO SECOND LAYER
FREQUENCY

C

AMPLITUDE
20 FIRST LAYER QUANTIZATION ERROR
FREQUENCY

D

AMPLITUDE
25 QUANTIZATION ERROR OF DECODING UP TO SECOND LAYER
FREQUENCY

FIG.4

A

30 21 ORIGINAL SOUND INPUT SIGNAL 11
f2 (SAMPLING FREQUENCY 24 kHz)
22, SAMPLE RATE CONVERTER

:

51 PERCEPTUAL WEIGHTED COEFFICIENT CALCULATION
35 SECTION

:

CODE C₁

24₁ FIRST CODER
(CELP CODER)
25₁ FIRST LOCAL DECODER
(CELP DECODER)
5 CODE C2
24₂ SECOND CODER
(CODER OF TRANSFORM CODING)
CODED CODE C
31 MULTIPLEXING CIRCUIT

10

B
FIRST FRAME
SECOND FRAME
THIRD FRAME

15 TIME

FIG.5

A

AMPLITUDE

20 FREQUENCY (kHz)

B

AMPLITUDE

23 Π POSITION INPUT SIGNAL25 12₁ DECODED SIGNAL

INPUT SIGNAL OF LOWER POSITION
FREQUENCY (kHz)

C

30 AMPLITUDE

29 INPUT SIGNAL OF HIGHER LAYER

13₁ QUANTIZATION ERROR OF LOWER LAYER
FREQUENCY (kHz)

35 D

AMPLITUDE

REFERENCE FOR QUANTIZATION WEIGHT USED FOR HIGHER

LAYER CODING
FREQUENCY (kHz)

E

5 AMPLITUDE
CODED/DECODED SIGNAL OF HIGHER LAYER
FREQUENCY (kHz)

F

10 AMPLITUDE
QUANTIZATION ERROR UP TO HIGHER LAYER
FREQUENCY (kHz)

FIG.6

15 22₁ SAMPLE RATE CONVERTER
36 ADAPTIVE CODEBOOK
37 NOISE CODEBOOK
CODE
GAIN
20 :
35 CONTROL SECTION
39 LINEAR PREDICTIVE ANALYSIS
41 QUANTIZATION
38 LINEAR PREDICTIVE SYNTHESIS FILTER
25 42 PERCEPTUAL WEIGHTED FILTER
43 FILTER COEFFICIENT CALCULATION SECTION

FIG.7

51 PERCEPTUAL WEIGHTED COEFFICIENT CALCULATION
30 SECTION
SPECTRUM ENVELOPE
47 PERCEPTUAL WEIGHTING SECTION
48 QUANTIZATION SECTION
46 LPC ANALYSIS
35 CODE
QUANTIZATION

FIG.8

56 DEMULTIPLEXING CIRCUIT

57₁ FIRST DECODER

(CELP DECODER)

5 57₂ SECOND DECODER

(DECODER OF TRANSFORM CODING)

DECODED SIGNAL OF LOWER LAYER

f₁ (SAMPLING FREQUENCY 16 kHz)

SAMPLE RATE CONVERTER

10 DECODING UP TO HIGHER LAYER

SIGNAL (SAMPLING FREQUENCY 24 kHz)

FIG.9

21 ORIGINAL SOUND INPUT SIGNAL

15 23₃ SAMPLE RATE CONVERTER

:

FIRST BAND SELECTION SECTION

:

SECOND BAND SELECTION SECTION

20 26₁ FIRST SAMPLE RATE CONVERTER26₂ SECOND SAMPLE RATE CONVERTER40₂ SECOND DECODING SECTION26₃ THIRD SAMPLE RATE CONVERTER40₃ THIRD DECODING SECTION25 72₄ PERCEPTUAL WEIGHTED COEFFICIENT CALCULATION
SECTION

:

24₁ FIRST CODER25₁ FIRST DECODER30 24₂ SECOND CODER25₂ SECOND DECODER24₃ THIRD CODER25₃ THIRD DECODER24₄ FOURTH CODER35 FIRST CODE C₁

FIRST DECODING SECTION

SECOND CODE C₂

THIRD CODE C_3 FOURTH CODE C_4

FIG.10

- 5 71₉ SAMPLE RATE CONVERTER
:
72₄ PERCEPTUAL WEIGHTED COEFFICIENT CALCULATION
SECTION
:
10 31 MULTIPLEXING CIRCUIT
21 ORIGINAL SOUND INPUT SIGNAL
66₂ SECOND BAND SELECTION SECTION
PASSBAND f
:
15 24₁ FIRST CODER
25₁ FIRST DECODER
26₁ FIRST SAMPLE RATE CONVERTER
31 MULTIPLEXING CIRCUIT
24₂ SECOND CODER
20 25₂ SECOND DECODER
26₂ SECOND SAMPLE RATE CONVERTER
24₃ THIRD CODER
25₃ THIRD DECODER
26₃ THIRD SAMPLE RATE CONVERTER
25 24₄ FOURTH CODER

FIG.11

- 56 CODE DEMULTIPLEXING CIRCUIT
57₁ FIRST DECODER
30 57₂ SECOND DECODER
57₃ THIRD DECODER
57₄ FOURTH DECODER
59₁ SAMPLE RATE CONVERTER
:
35